UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/663,390	09/15/2003	Dinei A. Florencio	MCS-032-03 (304924.01)	2920

27662 7590 08/18/2008
MICROSOFT CORPORATION
C/O LYON & HARR, LLP
300 ESPLANADE DRIVE
SUITE 800
OXNARD, CA 93036

EXAMINER	
LERNER, MARTIN	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
08/18/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/663,390	Applicant(s) FLORENCIO ET AL.	
	Examiner MARTIN LERNER	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *02 June 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 to 50* is/are pending in the application.

4a) Of the above claim(s) *22 to 50* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 to 17, 19, and 21* is/are rejected.

7)☒ Claim(s) *18 and 20* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Election/Restrictions

1. Applicants' election of Group I, Claims 1 to 21, in the reply filed on 10 December 2007 is acknowledged. Because Applicants did not distinctly and specifically point out the supposed errors in the restriction requirement, the election has been treated as an election without traverse (MPEP § 818.03(a)).

2. Claims 22 to 50 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected invention, there being no allowable generic or linking claim. Election was made **without** traverse in the reply filed on 10 December 2007.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1 and 8 to 11 are rejected under 35 U.S.C. 102(b) as being anticipated by *Shlomot*.

Regarding independent claim 1, *Shlomot* discloses a speech manipulation system for continuous speech playback over a packet network, comprising:

"storing data packets comprising a received audio data signal to a signal buffer" – Coded Speech Packets (CSPs) are received from packet network 100 at receiving speech terminal 204, and jitter buffer 260 acts as an intermediate buffer at the receiver end, allowing the packets to be played out of the jitter buffer 260 at regular intervals (column 4, lines 31 to 47: Figure 2);

"outputting parts of the signal present in the signal buffer as needed for signal playback" – jitter buffer 260 stores incoming speech packets before the packets are replayed; the stored packets can then be played out of the jitter buffer 260 at a regular predetermined replay rate (column 4, lines 46 to 51: Figure 2);

"analyzing the data packets contained in the signal buffer to determine whether any data packets are missing, having not been received into the signal buffer by an expected arrival time" – for voice data, packets that are lost or discarded result in gaps, silence, and clipping in real-time audio playback (column 1, lines 33 to 36); packets are analyzed to determine any of a normal event, where the time for incoming packets to the jitter buffer 260 is approximately equal to a predetermined standard replay rate, or a fast event when the rate of arrival of packets into the jitter buffer 260 is significantly higher than a predetermined replay rate, or a slow event when a rate of arrival between packets is significantly lower than the predetermined replay rate (column 8, lines 1 to 13: Figure 2); if packet P5 does not arrive at time t + 3 ("an expected arrival time"), a slow event occurs at time t + 3 (column 8, lines 18 to 22: Figure 4A);

"determining a maximum delay period for receiving any missing packets based on a current level of the signal buffer" – a slow event occurs when the rate of arrival

between packets into jitter buffer 260 is significantly lower than a predetermined replay rate, or is lower than a low threshold rate corresponding to a low threshold level (“based on a current level of the signal buffer”) of jitter buffer 260; thus, “a maximum delay period for receiving any missing packets” is defined by a slow event, which is a function of the fullness of jitter buffer 260;

“stretching at least part of the signal preceding the missing data packets present in the signal buffer, until any of receiving the missing data packets and exceeding the maximum delay period, when the analysis of the contents of signal buffer indicates that the length of the signal in the signal buffer is less than a predetermined threshold” – an underflow condition occurs when pointer 340 reaches a predetermined low level threshold (“when analysis of the contents of the signal buffer indicates that the length of the signal in the signal buffer is less than a predetermined threshold”) of jitter buffer 260 (column 6, lines 19 to 34: Figure 2); an underflow indicator from pointer 340 is used to signal an expansion function for expanding ("stretching at least part of the signal") a number of segments represented by a number of speech packets into a larger number of speech segments (column 7, lines 5 to 20: Figure 3); although P5 does not arrive at time t + 3, expansion logic 262 in speech decoder 240 expands packet P3 such that subsequent decoding results in speech packets S3A and S3B over two output speech segments (“at least part of the signal preceding the missing data”); packets P6 and P7 arrive late, but since P3 was already expanded, the buffer is not empty and P4 and P5 are played at a normal rate (column 8, lines 18 to 31: Figure 4A); thus, P5 is not

stretched because, although it is received late, it is subsequently received ("until any of receiving the missing data and exceeding the maximum delay period");

"compressing at least part of the signal present in the signal buffer when the analysis of the contents of the signal buffer indicates that the length of the signal in the signal buffer is greater than a predetermined threshold" – an overflow signal 266 is asserted only when pointer 340 is moved past a predetermined high level threshold ("greater than a predetermined threshold") of jitter buffer 260 (column 6, lines 19 to 25: Figure 2); if the jitter buffer in the time of arrival of the CSPs from the network exceeds a certain level, a jitter buffer can overflow; an overflow danger is detected when pointer 340 approaches an F location 330; an overflow indicator from pointer 340 is used to signal a compression function ("compressing at least part of the signal present in the signal buffer") for merging a number of stored speech packets into a smaller number of speech segments by speech decoder 240 (column 6, lines 37 to 47: Figure 3).

Regarding independent claim 8, *Shlomot* discloses a speech manipulation system for continuous playback over a packet network, further comprising:

"receiving and decoding data frames of an audio signal transmitted across a packet-based network" – CSPs are received from packet network 100 at receiving speech terminal 204, which includes a stripping unit 250 and a speech decoder 240 (column 4, lines 31 to 36: Figure 2); stripping unit 250 and speech decoder 240 perform functions of "decoding data frames of an audio signal;

"outputting one or more of the decoded frames present in the signal buffer when the analysis of the contents of the signal buffer indicates that the length of the signal in the buffer is between a predetermined minimum and a predetermined maximum buffer size" – if there is no jitter in the time of arrival of packets from network 100, buffer management unit 270 and fast/slow play unit 280 operate to pass the audio signal through the decoder path, and no compression or expansion is performed (column 7, lines 47 to 52: Figure 2); a normal event occurs where a time of arrival for incoming packets to jitter buffer 260 is approximately equal to a predetermined replay rate, and does not exceed a high threshold rate corresponding to a high threshold level or is lower than a low threshold rate corresponding to a low threshold level (column 8, lines 1 to 18: Figure 4A).

Regarding claim 9, *Shlomot* discloses that stored packets are played out of jitter buffer 260; a regular operation mode of a speech decoder would be to decode one CSP into a single speech segment of a predetermined length of 20 ms (column 4, lines 41 to 47: Figure 2); thus, a speech segment ("frame") is removed when it is played out.

Regarding claim 10, *Shlomot* discloses that voice packets may be lost (column 1, lines 33 to 36); packets may arrive at the right time for packets P3, P4, P9, P10, and P11, but packets may arrive late for packets P5, P6, P7, and P8 (column 7, lines 64 to 67: Figure 4A); an objective is to control playback to ensure that a listener will experience no discontinuity in speech (column 1, lines 55 to 59); thus, an objective is equivalent to "packet loss concealment" for "late loss packets".

Regarding claim 11, *Shlomot* discloses expanding a number of speech segments represented by a number of speech packets into a larger number of speech segments when an underflow indicator signals an expansion function (column 7, lines 5 to 8: Figure 3); a slow event occurs when an incoming rate of received packets is lower than a low threshold rate corresponding to a low threshold level in jitter buffer 260 (column 8, lines 9 to 18: Figure 3); thus, control is performed automatically according to the algorithm as a function of buffer content ("automatic control as a function of buffer content").

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 2 to 4, 12 to 15, and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Shlomot* in view of *Chong-White et al*.

Concerning claims 2, 12, and 13, *Shlomot* does not expressly provide for analyzing a contents of the signal buffer from a group including periodic content and aperiodic content, or voiced and unvoiced frames, prior to stretching decoded frames. However, *Chong-White et al*. teaches enhancing speech intelligibility using variable-rate time-scale modification, where vowel sounds (often referenced as voiced speech) and consonant sounds (often referenced as unvoiced speech) are processed from a buffer

702 so that some segments have lengthened time durations, corresponding to stretching, and other segments have compressed time durations, corresponding to compression. (Column 3, Lines 7 to 10: Figures 1 and 7) Specifically, formant transitions are emphasized through time expansion, and vowel segments are compressed. (Column 7, Lines 48 to 65: Figures 7 and 8) The objective is to enhance speech intelligibility due to consonant confusions in the presence of bandwidth reduction and packet loss. (Column 1, Line 60 to Column 2, Line 20) It would have been obvious to one having ordinary skill in the art to analyze contents of a signal buffer including periodic, aperiodic, voiced, and unvoiced segments prior to stretching or compressing as taught by *Chong-White et al*. in a speech manipulation system for continuous playback of *Shlomot* for a purpose of enhancing speech intelligibility in the presence of bandwidth reduction and packet loss.

Concerning claims 3 and 14, *Chong-White et al*. teaches stretching of segments involves searching using a cross-correlation to find a segment within a given tolerance ("identifying at least one segment . . . as a template") that has a maximum similarity ("exceeds a predetermined threshold") to the continuation of a last extracted segment (column 8, lines 39 to 43); the segment is matched with another segment using cross-correlation and waveform similarity criterion, and the segment and the best-matched segment are blended together by overlapping and adding the two segments together ("aligning and merging") (column 9, lines 18 to 41: Figure 8).

Concerning claims 4 and 15, *Chong-White et al*. teaches stretching consonants and unvoiced fricatives (column 7, lines 38 to 55), which are segments having "unvoiced" or "aperiodic" content, to increase speech intelligibility.

Concerning claim 19, *Chong-White et al*. teaches compressing a vowel following a consonant (column 7, lines 54 to 56), where a vowel is a "voiced frame"; procedures for stretching and compressing both involve searching using cross-correlation to find a segment having maximum similarity, and blending the best-matched segment together by overlapping and adding (column 7, lines 39 to 43; column 8, lines 18 to 41: Figure 8); one skilled in the art would know that the same procedure could be applied to "cutting out" matching signals for compressing and "inserting" matching signals for stretching.

7. Claims 7 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Shlomot* in view of *Hardwick et al*.

Shlomot discloses compression and expansion for lost speech packets, but omits compensating for clock drift. However, compensation for clock drift is known for receiving audio packets over a network. Specifically, *Hardwick et al*. teaches addressing a problem of a skew in time due to clock drift by compressing or expanding a data rate in a received signal as a function of the fullness of a buffer. (Column 9, Lines 7 to 45: Figure 5A) An objective is to minimize effects of mismatch between data rate states of two transceiver components in a signal transmission line. (Column 3, Lines 34 to 53) It would have been obvious to one having ordinary skill in the art to compensation for clock drift as taught by *Hardwick et al*. in a speech manipulation

system for continuous playback of *Shlomot* for a purpose of minimizing effects of mismatch between data rate states of two transceiver components.

8. Claims 5 to 6 and 16 to 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Shlomot* in view of *Chong-White et al*. as applied to claims 1, 2, 4, 8, and 15 above, and further in view of *Unno et al*.

Shlomot omits introducing a random rotation of the phase into frequency domain signals by applying at least one LPC filter to compute an LPC residual, computing at least one FFT from the LPC residual, introducing a random phase rotation into the coefficients, computing an inverse FFT, and applying an inverse LPC filter to the LPC residual to create at least one synthetic segment. However, *Unno et al*. teaches an enhancement to a mixed excitation linear predictive (MELP) coder, where one embodiment involves taking the Fourier magnitude of an LPC residual 23, introducing a random phase 64, performing an inverse DFT 93, and producing a mixed excitation signal 95 (column 11, lines 53 to 66: Figure 9). One skilled in the art would know that an LPC residual is then processed through an LPC synthesis filter to create synthesized speech. (Figures 2C and 8) An objective is to enhance the coded speech quality of a MELP coder for plosives (column 2, lines 31 to 56), which are unvoiced or aperiodic speech. It would have been obvious to one having ordinary skill in the art to perform the technique of random phase rotation of a frequency domain LPC residual as taught by *Unno et al*. in a speech manipulation system for continuous playback of *Shlomot* for a

purpose of enhancing the coded speech quality of plosives in an coder operating in accordance with MELP.

Response to Arguments

9. Applicants' arguments filed 02 June 2008 have been considered but are moot in view of the new grounds of rejection, necessitated by amendment.

Allowable Subject Matter

10. Claims 18 and 20 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Conclusion

11. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Chandos et al. discloses related prior art.

12. Applicants' amendment necessitated the new grounds of rejection presented in this Office Action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Martin Lerner/
Primary Examiner
Art Unit 2626
August 15, 2008